Risalah Kebijakan

Nomor 9, Oktober 2024



Kebijakan Pengembangan Bahasa:

Pengayaan Entri Kamus Besar Bahasa Indonesia



BADAN PENGEMBANGAN DAN PEMBINAAN BAHASA
KEMENTERIAN PENDIDIKAN DASAR DAN MENENGAH

badanbahasa.kemdikbud.go.id @BadanBahasa badanbahasakemendikbud Badan Pengembangan dan Pembinaan Bahasa @badanbahasa

Badan Bahasa
Bermartabat
Bermanfaat

RINGKASAN **EKSEKUTIF**

Badan Pengembangan dan Pembinaan Bahasa (Badan Bahasa) telah memprakarsai sebuah program nasional untuk memperluas dan meningkatkan kosakata bahasa Indonesia. Program ini bertujuan untuk menambahkan 80.000 entri baru ke dalam Kamus Besar Bahasa Indonesia (KBBI). Jumlah tersebut menambah jumlah total entri menjadi 200.000 entri.

Program telah dimulai dengan pelatihan peserta, pemilihan sumber data untuk menambah kosakata, dan pemanfaatan teknologi untuk mempercepat prosesnya. Lebih dari 100 orang dari berbagai unit pelaksana teknis (UPT) Badan Bahasa di seluruh Indonesia dilibatkan. Tim ini terdiri atas para ahli bahasa muda dan peserta magang yang sedang menjalani pelatihan leksikografi. Para ahli leksikografi senior yang berpengalaman bertugas juga menjadi mentor bagi tim junior. Lembaga riset di bidang korpus dan linguistik komputasi dilibatkan dalam bermitra untuk mengembangkan korpus bahasa Indonesia dan menggunakannya untuk pekerjaan leksikografi.

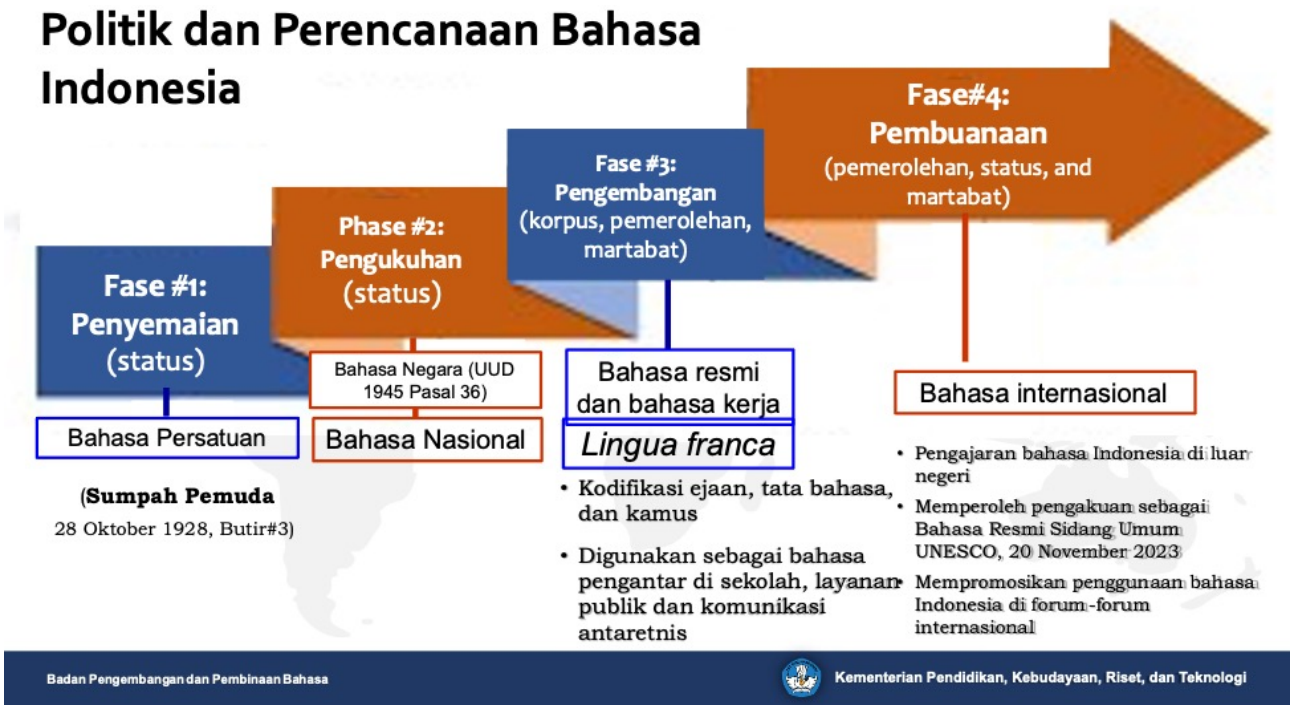
Salah satu alasan dibuatnya program ini adalah sebagai tindak lanjut dari pengakuan bahasa Indonesia sebagai bahasa resmi Sidang Umum UNESCO pada 20 November 2023. Berdasarkan hal itu, bahasa Indonesia harus meningkatkan kemampuan ekspresif dan kosakatanya. Dengan kata lain, pengayaan entri atau kosakata bahasa Indonesia yang terekam dalam KBBI merupakan sebuah keharusan.

1. PENDAHULUAN

1.1 Kebijakan dan Perencanaan Bahasa Indonesia

Salah satu proses dalam kebijakan dan perencanaan bahasa adalah perencanaan korpus yang berfokus pada pembakuan bentuk-bentuk bahasa, seperti ejaan, tata bahasa, dan kosakata. Perencanaan korpus ini merupakan tindak lanjut dari keputusan yang dibuat dalam perencanaan status yang mengidentifikasi dan mengakui varian bahasa yang memiliki status resmi dalam suatu masyarakat. Perencanaan status melegitimasi penggunaan bahasa ini oleh para penuturnya di berbagai ranah, termasuk kantor pemerintah, peradilan, sekolah, dan media massa.

Dari perspektif kebijakan dan perencanaan bahasa, bahasa Indonesia telah mengalami empat fase perkembangan yang berbeda. Fase-fase tersebut meliputi penyemaian, pengukuhan, pengembangan, dan pembuanaan. Fase penyemaian dimulai dengan pengenalan dan penggunaan awal istilah “bahasa Indonesia” yang kali pertama digunakan dalam Sumpah Pemuda 1928. Sebelumnya nama tersebut tidak ada. Bahasa Indonesia diakui sebagai pemersatu di antara para pemuda yang aktif dalam pergerakan kemerdekaan.



Gambar 1: Perjalanan Perumusan Kebijakan dan Perencanaan Bahasa Indonesia (Aziz, 2024)

Nama “bahasa Indonesia” diubah dari istilah “bahasa Melayu” yang berfungsi sebagai *lingua franca*, yaitu bahasa yang digunakan sebagai bahasa pergaulan dan digunakan secara meluas, terutama di kalangan pemuda, dalam transaksi perdagangan, dan komunikasi sehari-hari. Pada tingkat tertentu, bahasa pergaulan ini berfungsi sebagai media untuk memperkuat nasionalisme di antara para aktivis kemerdekaan. Nama “bahasa Indonesia” digunakan secara bersamaan untuk merujuk kepada tanah air dan bangsa, seperti yang diikrarkan dalam Sumpah Pemuda, yaitu *mengaku bertumpah darah yang satu, tanah air Indonesia; mengaku berbangsa yang satu, bangsa Indonesia; menjunjung bahasa persatuan, bahasa Indonesia*. Dalam konteks ini, bahasa Indonesia telah berubah atau bertambah fungsi, bukan semata-mata sebagai media komunikasi, melainkan juga telah menjadi ideologi baru dalam pergerakan kemerdekaan. Ideologi baru ini diperkenalkan kepada publik untuk menggantikan istilah bahasa Melayu yang sudah ada sebelumnya. Oleh karena itu, bahasa Indonesia dipromosikan ke publik sebagai faktor pendorong baru untuk menciptakan rasa kebersamaan.

Dalam UUD 1945 Pasal 36 dinyatakan bahwa “Bahasa negara adalah bahasa Indonesia”. Pasal ini tetap dipertahankan dalam bentuk aslinya, meskipun UUD 1945 telah diamandemen sebanyak empat kali. Pencantuman bahasa Indonesia sebagai bahasa negara dalam konstitusi ini menegaskan status bahasa Indonesia. Melalui konstitusionalisasi ini bahasa Indonesia ditetapkan sebagai bahasa negara yang memiliki kedudukan penting dalam pemerintahan negara, status yang berbeda dan memiliki gengsi lebih tinggi daripada bahasa daerah lain yang tersebar dan digunakan oleh masyarakat Indonesia di seluruh Indonesia.

Setelah tahap penetapan melalui Undang-Undang Dasar, norma dan standar bahasa Indonesia dikonseptualisasi. Pada fase ini, tata bahasa dikembangkan

sesuai dengan prinsip-prinsip tata bahasa modern. Kamus-kamus baru ditulis yang secara prinsip berbeda dengan kamus-kamus yang sudah ada dalam bahasa Melayu. Karya-karya sastra Indonesia juga berkembang cukup pesat dengan menggunakan tata bahasa Indonesia. Buku-buku pelajaran untuk sekolah-sekolah juga ditulis dalam bahasa Indonesia. Dengan didukung oleh instrumen-instrumen tersebut, yaitu tersedianya tata bahasa baku dan bahan ajar, bahasa Indonesia diajarkan secara formal di sekolah-sekolah sebagai mata pelajaran wajib dari tingkat pendidikan dasar hingga perguruan tinggi. Bahasa Indonesia juga digunakan di media massa, layanan publik, dan sebagai bahasa pengantar dalam pergaulan sehari-hari. Yang lebih menarik lagi, bahasa Indonesia mampu mengatasi masalah dan menjadi media komunikasi antaretnis di antara masyarakat Indonesia. Secara perlahan tetapi pasti, bahasa Indonesia menjadi identitas baru bagi masyarakat di negara yang baru saja merdeka ini.

Tahun 2009 menjadi saksi dinamika baru dalam politik bahasa Indonesia. Dengan persetujuan Dewan Perwakilan Rakyat, Pemerintah Indonesia memberlakukan undang-undang baru mengenai bendera, bahasa, dan lambang negara, serta lagu kebangsaan (UU Nomor 24 Tahun 2009). Undang-undang ini mengamanatkan pemerintah, antara lain, untuk mengembangkan strategi yang sistematis guna meningkatkan fungsi bahasa Indonesia menjadi bahasa internasional. Didorong oleh peraturan ini, pemerintah memulai internasionalisasi bahasa Indonesia secara lebih sistematis dan terorganisasi pada tahun 2015. Ini merupakan fase pembuanaan yang menjadi titik awal penyusunan konsep politik dan perencanaan pembuanaan bahasa Indonesia. Fase pembuanaan ini juga ditandai dengan pendirian pusat baru di Badan Pengembangan dan Pembinaan Bahasa untuk mengoordinasikan inisiatif ini di seluruh kementerian dan lembaga. Pusat-pusat studi dan pengajaran bahasa Indonesia di beberapa negara pun didirikan. Pengajaran bahasa Indonesia oleh perwakilan Indonesia di luar negeri, di sekolah-sekolah, universitas, dan mitra institusi lainnya ditawarkan sebagai bagian dari diplomasi budaya atau diplomasi lunak (*soft power diplomacy*). Kerja sama yang saling menguntungkan dengan para mitra ini terus berlanjut hingga saat ini. Melalui upaya diplomasi di UNESCO, usulan untuk menjadikan bahasa Indonesia sebagai bahasa resmi pada sidang umum UNESCO telah disetujui oleh Sidang Umum UNESCO pada tanggal 20 November 2023. Bahasa Indonesia menjadi bahasa resmi ke-10 dalam Sidang Umum UNESCO. Persetujuan ini menjadi tonggak sejarah baru bagi rencana globalisasi bahasa Indonesia.

Dengan berkaca pada langkah-langkah perencanaan korpus untuk bahasa Indonesia pada periode awal dan bahwa kodifikasi memainkan peran penting, masa pascapersetujuan UNESCO untuk bahasa Indonesia tersebut menghadapi dua sisi perencanaan. Di satu sisi, ada kodifikasi dan/atau standardisasi korpus yang akan digunakan oleh masyarakat Indonesia di dalam negeri, seperti dalam buku pelajaran di sekolah, media massa, perkantoran, dan perjanjian-perjanjian. Di sisi lain, ada pengembangan korpus yang ditargetkan untuk digunakan oleh pemelajar bahasa Indonesia sebagai bahasa asing. Model perencanaan bahasa yang linier-rekursif ini sangat khas dan dapat diterapkan pada bahasa Indonesia, terutama setelah bahasa Indonesia memperoleh statusnya di organisasi

internasional. Dengan model linier-rekursif ini perencanaan bahasa seolah kembali ke perencanaan status dan perencanaan pemerolehan/pembelajaran. Dengan kata lain, model perencanaan ini kembali memutar (rekursif) ke tahap sebelumnya, tetapi dengan kondisi sasaran yang berbeda.



Gambar 2: Model linier-rekursif kebijakan dan perencanaan bahasa Indonesia (Aziz, 2024)

Latar belakang budaya yang berbeda akan membutuhkan materi pembelajaran yang berbeda yang pada gilirannya akan membutuhkan korpus yang berbeda dengan strategi penyajian yang berbeda pula. Oleh karena itu, korpus yang saat ini sedang dikembangkan dan menjadi bagian dari entri di KBBI juga mempertimbangkan konteks ini. Program penambahan 80.000 entri baru ke dalam KBBI pada tahun 2024 hingga mencapai 200.000 entri direncanakan untuk memenuhi kebutuhan dalam menyediakan pilihan kosakata bahasa Indonesia yang lebih luas bagi para pengguna bahasa Indonesia.

1.2 Gambaran Umum Program Nasional Badan Pengembangan dan Pembinaan Bahasa

Program nasional yang diberi nama “Pengayaan Entri atau Kosakata Bahasa Indonesia” ini merupakan inisiatif pertama yang menetapkan target entri dengan skala besar. Kelayakan program ini didukung oleh estimasi Badan Pengembangan dan Pembinaan Bahasa tentang sumber daya kebahasaan yang dapat dimobilisasi untuk usaha ini. Kekuatan sumber daya tersebut terletak pada ketersediaan jumlah sumber data yang terdiri atas bahasa daerah, yang mencakup kata-kata budaya atau kata-kata dengan konsep unik, istilah-istilah padanan dalam bidang ilmu pengetahuan yang merupakan hasil pengindonesiaan istilah, kata-kata pinjaman dari bahasa asing yang terus membanjiri bahasa Indonesia,

kata-kata atau ungkapan dalam bahasa Indonesia percakapan atau informal, dan bentuk-bentuk kreativitas bahasa lainnya. Melimpahnya data kebahasaan ini merupakan salah satu sumber pengayaan kosakata bahasa Indonesia yang dapat dengan mudah diakses dan siap diolah secara digital.

Sumber daya manusia yang terlibat dalam program ini mencakup lebih dari 100 orang ahli leksikografi, yang berasal dari unsur internal dan eksternal Badan Pengembangan dan Pembinaan Bahasa. Kedua kelompok tersebut bekerja sama untuk mencapai tujuan yang telah ditetapkan. Selain itu, kelompok editorial ditambah dengan kelompok kontributor pemula yang berasal dari mahasiswa jurusan bahasa dari 18 universitas yang berbeda. Sebanyak 180 kontributor mahasiswa terlibat dalam program ini, dengan 10 mahasiswa dari masing-masing 18 universitas yang berpartisipasi. Berbeda dengan pola rekrutmen tim editorial eksternal, kontributor dari kelompok mahasiswa ini membantu Badan Pengembangan dan Pembinaan Bahasa melalui platform magang mahasiswa. Dengan demikian, mereka mendapatkan penilaian untuk mata kuliah magang. Setiap mahasiswa ditugaskan untuk membuat 50 entri baru yang hal tersebut berarti program ini menghasilkan 9.000 entri baru pada akhir program. Sebelum melaksanakan tugas, mahasiswa menerima pembekalan selama tiga hari, dan kemajuan mereka diawasi oleh dosen pembimbingnya.

1.3 Tujuan: Memperluas dan Meningkatkan Kosakata Bahasa Indonesia

Tujuan program ini adalah untuk meningkatkan jumlah entri dalam KBBI agar bisa mencapai 200.000, yang mencakup penambahan kata turunan, kata majemuk, peribahasa, idiom, dan bentuk-bentuk lain yang mencerminkan perkembangan bahasa dan masyarakat. Meskipun mungkin tampak ambisius, angka ini memiliki dampak psikologis yang signifikan terhadap bahasa Indonesia, yang selama lebih dari tiga puluh tahun hanya mampu mencatatkan 120.000 entri dalam KBBI. Jumlah ini juga dapat menunjukkan betapa kayanya kosakata yang dibutuhkan oleh sebuah bahasa untuk menjadi bahasa yang lebih modern dan digunakan secara luas, serta meningkatkan status bahasa Indonesia sebagai bahasa internasional. Beberapa pemikiran berikut ini menjadi pertimbangan .

Pertama, fleksibilitas dan kemampuan beradaptasi adalah faktor kunci. Bahasa dengan kosakata yang luas cenderung lebih fleksibel dan mudah beradaptasi dengan konteks dan situasi yang beragam. Hal ini membuat bahasa Indonesia menjadi bahasa yang lebih mudah diakses di berbagai bidang, termasuk ilmu pengetahuan, teknologi, bisnis, dan seni.

Kedua, variasi ekspresi yang diberikan oleh kosakata yang luas memungkinkan penutur untuk menyampaikan ide, konsep, dan nuansa makna dengan lebih tepat dan spesifik. Selain itu, jumlah kosakata yang besar memungkinkan ekspresi yang akurat dan tepat untuk konsep-konsep yang kompleks yang sangat penting dalam wacana akademis dan profesional.

Ketiga, pengembangan kosakata sebagai bentuk penerimaan dan pengembangan. Bahasa dengan kosakata yang besar lebih mudah diterima dan dikembangkan di berbagai komunitas. Penutur bahasa akan lebih cenderung mengadopsi bahasa dengan kosakata besar yang relevan dengan kebutuhan mereka.

1.4 Pentingnya Program Pengayaan Kosakata Bahasa Indonesia

Perencanaan korpus dan program pengayaan kosakata saling berkaitan erat, yang keduanya merupakan komponen kunci dalam pengembangan secara sistematis dan upaya modernisasi bahasa. Hubungan antara keduanya dapat dilihat dengan cara berikut.

1. Merefleksikan perubahan masyarakat
   Seiring dengan perkembangan masyarakat, konsep baru, teknologi, dan pengaruh global mengharuskan terciptanya kosakata baru. Perencanaan korpus memastikan bahwa kata-kata baru ini diintegrasikan secara sistematis ke dalam bahasa, sementara program pengayaan kosakata menyediakan mekanisme untuk pengenalan dan penyebarannya.
2. Standardisasi dan aksesibilitas
   Perencanaan korpus memainkan peran penting dalam menstandarkan bahasa yang sangat penting untuk komunikasi dan pendidikan. Program pengayaan kosakata berkontribusi dalam memastikan bahwa entri baru dapat diakses oleh semua penutur dan pelajar bahasa yang dapat mendukung kesatuan bahasa.
3. Pelestarian dan promosi budaya
   Baik perencanaan korpus maupun program pengayaan kosakata dapat digunakan untuk melestarikan dan mempromosikan warisan budaya Indonesia. Dengan mendokumentasikan dan membakukan dialek, idiom, dan ungkapan daerah, program ini turut membantu menjaga keragaman budaya dalam bahasa nasional.
4. Pengembangan pendidikan
   Perencanaan korpus menyediakan kerangka kerja untuk materi pendidikan dan memastikan konsistensi serta kejelasan dalam pengajaran bahasa. Program pengayaan kosakata mendukung hal ini dengan memperbarui materi pendidikan dengan kata-kata dan makna baru, sehingga menjaga kurikulum tetap relevan dan sesuai dengan perkembangan zaman.
5. Penelitian linguistik
   Perencanaan korpus mendukung penelitian linguistik dengan menyediakan data bahasa yang terstruktur dan terstandardisasi. Program pengayaan kosakata memperkaya data ini dengan bentuk-bentuk linguistik baru, sehingga memberikan sumber daya yang dinamis dan komprehensif bagi para peneliti untuk mempelajari perubahan dan perkembangan bahasa.

2. KONDISI SAAT INI

Salah satu metode yang paling umum digunakan dalam penyusunan KBBI adalah program membaca yang selama ini hasilnya relatif lambat. Sebelum edisi kelima, semua pemutakhiran KBBI dilakukan secara manual di atas kertas. Pendekatan ini menghasilkan jumlah entri tambahan yang terbatas karena terbatasnya jumlah editor yang terlibat dalam proses tersebut. Tabel berikut ini memberikan gambaran umum tentang sejarah perkembangan KBBI.

Tabel 1. Pertumbuhan Entri KBBI per Edisi

Edisi	Tahun	Jumlah Entri
KBBI I	1988	62.000
KBBI II	1991	72.000
KBBI III	2001	78.000
KBBI IV	2008	90.000
KBBI V	2016	112.000
KBBI VI	2023	120.000

KBBI telah mengalami dua periode pemutakhiran: periode manual atau berbasis kertas yang berlangsung selama KBBI I hingga KBBI IV dan periode digital yang dimulai dari KBBI V hingga KBBI VI. Selama periode manual, sekitar 28.000 entri ditambahkan atau penambahan rata-rata 1.400 entri baru per tahun selama rentang waktu 20 tahun. Dari tahun 2016 hingga 2023, KBBI mengalami periode ekspansi digital dengan tambahan 8.000 entri selama kurun waktu tujuh tahun. Laju penambahan entri selama periode manual dan digital tidak berbeda secara signifikan. Tingkat rata-rata penambahan entri selama periode digital bahkan lebih rendah daripada periode manual, yaitu sekitar 1.140 entri per tahun. Situasi ini kurang akomodatif, terutama mengingat bahwa bahasa Indonesia adalah bahasa yang sedang berkembang dan sedang mengalami proses perluasan kosakata yang dinamis.

KBBI Daring saat ini merupakan aplikasi kamus daring yang paling banyak digunakan untuk pencarian. Menurut statistik yang tersedia di situs web KBBI (https://kbbi.kemdikbud.go.id/Beranda/Statistik), tingkat pencarian harian saat ini mendekati 100.000 pencarian. Pada 19 Januari 2021, misalnya, KBBI mencapai 100 juta pencarian, menandai hari ke-1.543 setelah diluncurkan. Tren ini diperkirakan akan terus berlanjut dengan kecepatan yang stabil, sehingga menjadikan KBBI sebagai referensi bahasa Indonesia yang paling otoritatif.

3. TANTANGAN DAN KETERBATASAN

Tantangan dan keterbatasan yang dihadapi oleh Badan Pengembangan dan Pembinaan Bahasa dapat diklasifikasikan ke dalam tiga kategori. Yang pertama mencakup kendala terkait jumlah dan variasi sumber data dan referensi. Yang kedua adalah keterbatasan ahli leksikografi. Kategori ketiga mencakup tantangan dalam mengadopsi teknologi untuk penyusunan kamus.

Terlepas dari melimpahnya sumber data mentah untuk pengayaan kosakata bahasa Indonesia, tantangan yang signifikan muncul dari kurangnya format digital yang tersedia, terutama untuk data dari bahasa daerah. Bahkan setelah pengumpulan data lapangan, data tersebut masih dalam bentuk mentah, sehingga perlu didigitalisasi. Pangkalan Data Kamus KBBI Daring, sebuah aplikasi pangkalan data kamus, menyediakan platform untuk memasukkan calon entri yang akan diusulkan. Namun, proses penginputan data sangat kompleks, membutuhkan program yang dapat menangani berbagai jenis data secara efektif. Untuk mengatasi masalah tersebut, diadakan pelatihan penginputan data, walaupun masih terdapat anggota tim penginput yang masih mengalami kesulitan.

Tantangan kedua yang dihadapi adalah terbatasnya jumlah editor dan tingkat keahlian yang berbeda-beda. Dapat dimaklumi bahwa tidak semua pegawai di Badan Pengembangan dan Pembinaan Bahasa terlibat dalam pekerjaan leksikografi. Dalam pembentukan tim, terdapat beberapa anggota yang belum berpengalaman dalam bidang leksikografi. Hal tersebut menyebabkan perbedaan kemampuan ketika bekerja bersama anggota tim lainnya. Selain itu, fokus mereka pada leksikografi sering kali terganggu oleh tanggung jawab pekerjaan yang lain.

Kemampuan staf untuk secara efektif menguasai dan memanfaatkan teknologi baru merupakan hal yang menjadi perhatian penting. Selama masa transisi dari KBBI I ke KBBI IV, para editor terbiasa menggunakan metode manual berbasis kertas. Namun, dengan dirilisnya KBBI V, seluruh proses editorial menjadi digital dan bergantung pada aplikasi tertentu. Beberapa ahli leksikografi yang berpengalaman menghadapi tantangan dalam beradaptasi dengan platform teknologi baru ini yang mengakibatkan masa penyesuaian yang lama bagi tim editor KBBI. Fase transisi yang berkepanjangan ini, sampai batas tertentu, menghambat produktivitas tim secara keseluruhan.

4. SOLUSI DAN STRATEGI IMPLEMENTASI

Penambahan entri KBBI menargetkan sejumlah besar entri baru. Proses ini membutuhkan strategi implementasi yang efisien, efektif, dan berhasil guna. Berikut ini adalah beberapa strategi implementasi yang disusun untuk mencapai target tersebut.

4.1 Lokakarya dan Pelatihan Leksikografi

Untuk memastikan bahwa kegiatan ini terfokus, setidaknya 16 lokakarya leksikografi yang masing-masing berlangsung selama 5 hari diselenggarakan untuk 40—50 tim editorial secara bergilir. Dengan memusatkan tim editorial di lokasi tertentu yang kondusif, diharapkan ada kemudahan untuk mencapai tujuan bagi setiap orang dalam setiap lokakarya. Sebagai contoh, tim validator harus memvalidasi hampir 400 entri baru dalam setiap lokakarya. Pada saat tulisan ini dibuat, 14 lokakarya telah diselenggarakan dan lebih dari 55.000 entri baru telah divalidasi (per 20 Oktober 2024) . Kegiatan ini diulang hingga 15 kali,

dan setiap lokakarya memiliki laporan berkala yang berfungsi sebagai alat kontrol pencapaian target. Mereka yang tidak mencapai target selama lokakarya harus mencapainya pada lokakarya berikutnya.

Pelatihan leksikografi terutama dirancang untuk editor eksternal dan juga untuk editor junior yang memiliki pengalaman terbatas. Pelatihan diberikan dalam beberapa format yang berbeda. Program pelatihan awal dilakukan dalam bentuk kursus tiga hari untuk kontributor mahasiswa yang dilaksanakan dalam tiga gelombang terpisah. Materi pelatihan mencakup pengenalan aplikasi KBBI Daring, termasuk gambaran umum fitur-fiturnya, tinjauan umum penggunaan korpus untuk leksikografi, pembentukan kata dalam morfologi bahasa Indonesia, dan latihan praktis untuk mengusulkan entri dengan aplikasi KBBI Daring. Mengingat tujuan dari pelatihan ini adalah untuk mengusulkan atau berkontribusi pada entri baru, materi pelatihan pun disesuaikan. Misalnya, materi pembentukan kata dalam morfologi bahasa Indonesia relevan dengan jenis usulan entri yang diharapkan di KBBI Daring, yaitu kata turunan berimbuhan.

Jenis pelatihan lainnya adalah untuk para penyunting dengan tugas utama menyunting entri. Pelatihan ini diberikan beberapa kali selama penyunting menjalankan tugasnya. Hal ini lebih tepat disebut sebagai pendampingan karena setiap editor senior menjadi pendamping bagi beberapa editor junior. Transfer pengetahuan dan pengalaman ini membuat lokakarya leksikografi menjadi sangat dinamis. Beberapa keputusan dan kesepakatan penting dihasilkan dari lokakarya ini.

4.2 Sidang Komisi Bahasa Daerah

Para leksikografer senior juga memberikan bimbingan kepada tim editor di Unit Pelaksana Teknis di daerah. Tujuannya adalah untuk memberikan pengarahan kepada tim redaksi lokal, yang tidak diundang ke lokakarya. Kegiatan ini dikemas dalam sebuah kegiatan yang disebut Sidang Komisi Bahasa Daerah (SKBD). Dalam SKBD ini, para editor senior mengajari para editor junior bagaimana memilih dan memilah entri potensial dari bahasa daerah yang berpotensi menjadi neologisme dalam bahasa Indonesia. Adapun pemilihan kosakata bahasa daerah harus memenuhi syarat sebagai berikut.

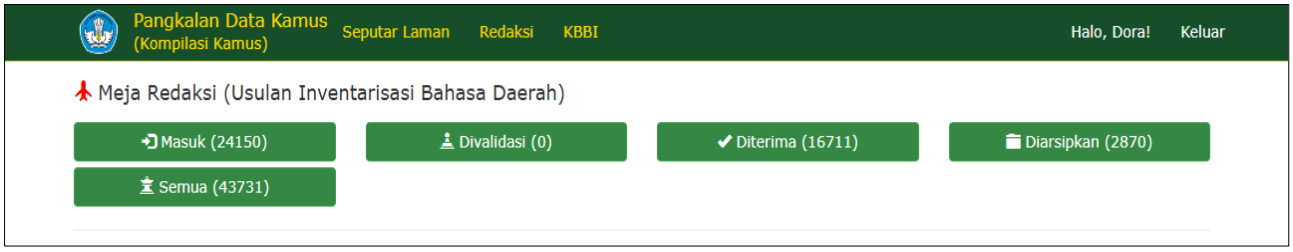
1. Kata tersebut berkonsep unik dan tidak ada padanannya dalam dalam bahasa Indonesia.
2. Secara bunyi, kata baru tersebut harus eufonik atau sesuai dengan sistem bunyi bahasa Indonesia.
3. Kata tersebut dapat dibentuk sesuai dengan kaidah morfologi atau pembentukan kata dalam bahasa Indonesia.
4. Kata tidak berkonotasi negatif atau mengandung isu SARA.
5. Kata tersebut berfrekuensi tinggi atau banyak digunakan dalam konteks penggunaan bahasa Indonesia secara umum.

SKBD ini sudah dilakukan sejak beberapa tahun lalu dan pada tahun ini lebih dioptimalkan karena ada target yang ingin dicapai. Dalam beberapa tahun

terakhir, SKBD telah memberikan data 24.150 calon entri, yang disimpan dalam bentuk digital dalam pangkalan data kamus, yang berada dalam satu platform yang sama dengan KBBI. Pangkalan data kamus merupakan aplikasi untuk menyimpan data calon entri yang diusulkan untuk dimasukkan ke dalam KBBI, tidak hanya dari hasil inventarisasi bahasa daerah, tetapi juga dari kamus bidang ilmu dan kamus dwibahasa.

Data yang dimasukkan ke dalam aplikasi pangkalan data kamus juga harus melalui beberapa proses penyaringan atau seleksi. Terdapat tiga kategori atau status calon entri, yaitu kategori *masuk* untuk calon entri yang sudah dimasukkan ke dalam aplikasi, *divalidasi* untuk entri yang sudah dikoreksi dan siap diusulkan, serta status *diterima* untuk entri yang sudah divalidasi dan siap diterima sebagai usulan entri baru dalam KBBI.

Tabel berikut menunjukkan hasil inventarisasi kosakata bahasa daerah pada tahun 2021. Angka 24.150 menunjukkan jumlah entri yang telah dimasukkan ke dalam aplikasi. Dari jumlah tersebut, terdapat 16.711 entri yang telah diseleksi dan siap dimasukkan sebagai usulan entri dalam KBBI, dan 2.870 entri yang diarsipkan karena tidak lolos seleksi.



Gambar 3. Jumlah entri potensial dari bahasa lokal

4.3 Pemilihan Sumber Data untuk Penambahan Kosakata

Aplikasi KBBI Daring dilengkapi dengan Pangkalan Data Kamus, yang merupakan tempat penyimpanan entri-entri baru yang potensial yang diperoleh dari berbagai sumber. Sumber pertama adalah kamus khusus yang telah disusun sebelumnya. Meskipun beberapa entri berupa istilah bidang ilmu sudah masuk dalam KBBI, belum semua isi kamus tersebut masuk dalam KBBI. Dalam daftar inventaris terdapat setidaknya 150 kamus yang telah disusun.

Sumber lain yang potensial untuk pengayaan entri adalah hasil inventarisasi kosakata bahasa daerah. Pada prinsipnya, inventarisasi kosakata bahasa daerah ini merupakan daftar yang diperoleh dari kegiatan SKBD. Sidang ini diadakan secara rutin di pusat atau UPT Badan Bahasa.

Sumber daya lain yang telah menjadi kegiatan rutin adalah pencarian data melalui program membaca bagi tim perkamusan. Hal ini sangat bergantung pada kejelian dan ketekunan para pekamus serta menjadi pilihan para pekamus senior yang tidak terbiasa bekerja secara digital.

Korpus Indonesia (KOIN) adalah salah satu sumber data yang tersedia secara terbuka. Namun, saat ini ada masalah dengan aksesnya yang tidak mudah

karena masalah server yang tidak memadai. Selain itu, jumlah data KOIN yang tersedia saat ini baru mencapai 25 juta kata, yang masih jauh dari cukup untuk sebuah korpus leksikografi. Karena keterbatasan ini, aplikasi KOIN menjadi aplikasi yang paling jarang digunakan untuk tujuan leksikografi. Sumber korpus yang memadai adalah *LCC Indonesian 2022: powered by CQPweb* (lancs.ac.uk) dengan 232.853.664 kata yang telah dilematisasi. Sumber ini terbuka dan oleh karena itu dapat diajarkan secara gratis kepada mahasiswa magang selama kegiatan pembekalan dan kepada editor eksternal. Sumber korpus berbayar lainnya yang hanya digunakan oleh pekamus internal Badan Pengembangan dan Pembinaan Bahasa adalah *Indonesian Web 2020* (*idTenTen20*), sebuah korpus bahasa Indonesia untuk keperluan umum yang mencakup berbagai genre, topik, jenis teks, dan sumber web. Direkomendasikan untuk penggunaan umum dan khusus. Data pada sumber ini sudah dianotasi oleh TreeTagger. Data diunduh pada bulan Juni—Agustus 2020 dengan total 4.432.864.160 kata dari 12.804.786 dokumen.

5. KOMPOSISI DAN PERAN TIM

KBBI adalah salah satu produk yang paling menonjol dan telah menjadi rujukan paling otoritatif sejak pertama kali diterbitkan pada tahun 1988. Sejak saat itu, KBBI diperbaharui hampir setiap 5 tahun saat Kongres Bahasa Indonesia diselenggarakan. Selama ini penambahan entri pada setiap edisi tidak terlalu besar (lihat bagian perkembangan KBBI). Oleh karena itu, sejak awal rencana program pada tahun 2023, Badan Pengembangan dan Pembinaan Bahasa telah mengembangkan skema penambahan tim editorial yang terdiri atas dua elemen: internal dan eksternal. Penambahan editor internal melibatkan tim editor dari seluruh balai/kantor bahasa (UPT) di 30 provinsi di Indonesia. Keikutsertaan mereka dimungkinkan karena mereka tergabung dalam Kelompok Kepakaran dan Layanan Profesional (KKLP) yang sama, khususnya di bidang perkamusan dan peristilahan.

Keterlibatan lebih dari 100 penyunting baru tersebut pada awalnya adalah untuk mencari calon entri dari bahasa daerah yang mereka kuasai untuk menjadi entri dalam KBBI. Dalam program besar ini, peran mereka bergeser menjadi editor umum, yang juga memproses usulan entri, tidak hanya dari bahasa daerah. Kategori setiap editor dari UPT ini juga tidak sama. Anggota tim yang lebih berpengalaman berperan sebagai penyelia bagi anggota tim yang kurang berpengalaman.

Untuk menjaga kualitas, penyuntingan dilakukan secara berjenjang. Tim editorial KBBI dibagi menjadi 3 tingkatan, yaitu (dari yang terendah ke yang tertinggi): editor, redaktur, dan validator. Komposis jumlah dan susunannya adalah sebagai berikut.

- editor 165
- redaktur 46
- validator 15

Dengan komposisi ini, setiap satu usulan entri baru harus melalui tahap seleksi dan penyuntingan secara ketat dan berjenjang.

Selain tim editor internal dari Badan Bahasa, 20 editor eksternal direkrut untuk memberikan dukungan bagi program ini. Dari 608 pelamar, 20 orang dipilih untuk posisi tersebut untuk bekerja dengan masa 10 bulan. Kriteria penilaian untuk editor eksternal ini ada tiga: portofolio atau pengalaman, pendidikan, dan ketersediaan mereka untuk posisi lain.

6. DUKUNGAN TEKNOLOGI

Pelibatan teknologi dalam program ini merupakan suatu keniscayaan. Untuk mencapai tujuan sebesar itu, metode-metode konvensional yang biasa digunakan selama ini tidak tepat diimplementasikan. Selama ini pelibatan teknologi bukanlah hal yang baru, meskipun masih dalam porsi yang masih terbatas dengan platform digital dalam proses penyuntingan.

6.1 Kolaborasi dengan Mitra Lembaga Riset Pemrosesan Data

Kolaborasi dengan mitra lembaga riset dalam program pengayaan kosakata di Indonesia ini memainkan peran penting dalam meningkatkan efektivitas dan jangkauan inisiatif ini. Berikut adalah bagaimana kolaborasi tersebut berkontribusi pada pengayaan kosakata bahasa Indonesia.

a. Berbagi keahlian: lembaga riset ini membawa pengetahuan khusus dalam linguistik dan pendidikan bahasa, yang sangat berharga dalam mengidentifikasi kata-kata yang paling relevan dan berdampak untuk ditambahkan ke daftar kosakata.
b. Analisis data: lembaga riset ini dapat memanfaatkan analisis data tingkat lanjut untuk mempelajari pola penggunaan bahasa, membantu menemukan celah dalam kosakata yang ada dan menyarankan entri baru yang mencerminkan tren linguistik terkini.
c. Integrasi teknologi: lembaga riset ini dapat memperkenalkan teknologi mutakhir, seperti pembelajaran mesin dan pemrosesan bahasa alami, untuk menyederhanakan proses pemilihan dan integrasi kosakata.

6.2 Perluasan Korpus Bahasa Indonesia

Perluasan korpus bahasa Indonesia merupakan upaya komprehensif yang melibatkan kolaborasi teknologi, akademisi, dan publik. Hal tersebut adalah bukti komitmen untuk mempertahankan bahasa Indonesia sebagai bahasa yang hidup dan berkembang seiring dengan kebutuhan penuturnya dan pergeseran budaya bangsa. Beberapa inisiatif dan metodologi baru digunakan untuk pengembangan dan modernisasi bahasa Indonesia.

Saat ini sudah ada korpus yang cukup besar (mengandung 3,7 miliar kata) yang digunakan dan akan diperluas menjadi dua kali lipat ukurannya. Alasan utamanya

adalah untuk meningkatkan jumlah kata dasar baru yang potensial. Selain itu, ada dua alasan utama lain yang melatarbelakangi pemerluasan korpus, yaitu:

a. untuk tetap mengikuti perkembangan: korpus yang ada saat ini berasal dari tahun 2020, yaitu berusia 4 tahun dan belum mencakupi kosakata baru yang penting (misalnya kata-kata baru terkait pandemi) dan
b. menambah ruang lingkup dan cakupan korpus: contoh kalimat dan kolokasi akan memiliki kualitas yang lebih baik dengan lebih banyak data yang beragam dan luas cakupannya. Prinsip "makin besar, makin baik" berlaku dalam hal ini.

6.3 Penggunaan *Dictionary Writing System* (DWS)

Seperti yang telah disebutkan sebelumnya, keterlibatan mitra dari lembaga riset dan pengolah data menjadi bagian penting dari tahap ini. Mitra ini membantu mempercepat pengerjaan entri baru dengan menggunakan teknologi yang disebut Dictionary Express. Pada tahap awal, diadakan uji coba dengan contoh data (100—200 entri lengkap) untuk melihat struktur data kamus dalam format sumber atau sebagai hasil ekspor ke XML atau sejenisnya. Hal tersebut dilakukan untuk memastikan komponen apa saja yang diperlukan dan biasanya ada dalam setiap entri agar dapat dihasilkan data yang semirip mungkin dengan data KBBI yang sebenarnya.

Proses percepatan pengayaan KBBI bersifat iteratif dan setiap langkah memerlukan campur tangan manusia dari para pekamus. Proses campur tangan manusia meliputi hal-hal berikut.

1) pemilihan calon lema yang valid dari daftar korpus. Proses ini dilakukan menurut tahapan sebagai berikut.
   a) memilih calon lema yang valid
   b) menyunting lema dari kesalahan yang dibuat pada proses lematisasi dan anotasi otomatis
2) penetapan makna pada lema yang dipilih
3) seleksi contoh korpus yang diidentifikasi pada langkah sebelumnya.

Sebelum fase intervensi manusia, mitra dari lembaga riset harus menyiapkan sumber data. Salah satu hal penting adalah membuat korpus bahasa Indonesia baru berukuran miliaran yang terkini sebagai sumber dasar untuk entri kamus. Hasil yang diharapkan dapat berupa hal-hal sebagai berikut:

1) model makna kata berdasarkan penempatan kata khusus;
2) algoritma contoh kamus yang baik (GDEX);
3) ukuran kekuatan kolokasi; dan
4) ukuran kesamaan kata (untuk sinonim, antonim, dan kata-kata yang mirip).

Tahap selanjutnya adalah penyuntingan akhir yang bertujuan untuk menyatukan semua informasi untuk setiap lema. Tugas pekamus adalah meninjaunya

secara keseluruhan dan mengoreksi kemungkinan kesalahan dari fase anotasi sebelumnya.

7. REKOMENDASI

Berikut adalah beberapa rekomendasi yang menyoroti pentingnya program pengayaan entri KBBI.

1) Pelestarian dan pengembangan linguistik
   Program ini harus memastikan bahwa kekayaan bahasa Indonesia dilestarikan dan terus berkembang dengan menggabungkan kata-kata dan ekspresi baru yang muncul dari berbagai bidang kehidupan, termasuk teknologi, sains, dan budaya populer.
2) Sumber daya pendidikan
   Program ini harus berfungsi sebagai sumber daya pendidikan yang sangat berharga bagi siswa, pendidik, dan peneliti, yang menyediakan referensi yang dapat dipercaya untuk memahami kosakata bahasa Indonesia yang terus berkembang.
3) Refleksi budaya
   Bahasa adalah refleksi budaya, dan pengayaan entri KBBI harus dapat membantu untuk mendokumentasikan dan mempromosikan aspek budaya Indonesia yang unik, termasuk dialek, peribahasa, dan idiom setempat.
4) Peningkatan komunikasi
   Dengan memperluas leksikon, program ini harus membantu untuk meningkatkan komunikasi di antara penutur bahasa Indonesia, sehingga memungkinkan percakapan yang lebih tepat dan ekspresif.
5) Dukungan bagi pembelajar bahasa
   Bagi penutur non-asli dan pemelajar bahasa Indonesia, daftar kosakata yang diperkaya menawarkan alat yang komprehensif untuk meningkatkan kemahiran bahasa dan pemahaman makna yang bernuansa.
6) Standardisasi
   Program ini harus berkontribusi pada standardisasi bahasa Indonesia, yang penting untuk dokumentasi resmi, media, dan materi pendidikan.
7) Program pengayaan awal KBBI harus dapat menjadi bukti komitmen Indonesia untuk membina lingkungan linguistik yang dinamis dan kuat yang tidak hanya mengakomodasi, tetapi juga merayakan keragaman dan kemampuan beradaptasi bahasa Indonesia.

Singkatnya, perencanaan korpus dan program pengayaan kosakata harus dapat mejadi proses simbiosis yang mendukung pertumbuhan dan dinamika bahasa Indonesia, memastikan bahwa bahasa Indonesia tetap menjadi media komunikasi, pendidikan, dan ekspresi budaya yang kuat di Indonesia.

DAFTAR PUSTAKA

Aziz, E. Aminudin. 2024. "Language Policy and Planning in Indonesia: Local Language Preservation Policy and National Language Globalization Initiatives" Paper presented in The 21st AILA World Congress, Malaysia, 11-16 August 2024.

Baldauf, R.B. Jr. 2008. "Language Planning: Corpus Planning" in *Annual Review of Applied Linguistics*, Volume 10, March 1989. DOI: https://doi.org/10.1017/S0267190500001173

Kamajaya, I., Moeljadi, D., and Amalia, D. 2017. KBBI Daring: A Revolution in The Indonesian Lexicography in Kozem, I. (ed.) *Electronic lexicography in the 21st century*. Leiden: Proceedings of e-Lex 2017 conference.

Luthfita, W., and Yanita, Selly, S.R. 2023. "Digitalizing Local Language Dictionary: Challenges and Opportunities" in Amalia, D. (ed.) *Lexicography*

Risalah Kebijakan
Nomor 9, Oktober 2024

BADAN PENGEMBANGAN DAN PEMBINAAN BAHASA
KEMENTERIAN PENDIDIKAN DASAR DAN MENENGAH

Pengarah:
E. Aminudin Aziz

Penyelia:
Imam Budi Utomo

Penyusun:
E. Aminudin Aziz
Dora Amalia

Desain Grafis:
Nurjaman

Diterbitkan oleh:
Badan Pengembangan dan Pembinaan Bahasa
Jalan Daksinapati Barat IV, Rawamangun, Jakarta Timur

